Limit theorems for Multilevel estimators with and without weights

Daphné Giorgi, Vincent Lemaire, Gilles Pagès
LPMA - UPMC

10 January 2017

General framework

- $Y_0 \in \mathbf{L}^2(\mathbf{P})$ non-simulatable real random variable.
- Aim: compute $I_0 = \mathbf{E}\left[Y_0\right]$
  - given an accuracy $\varepsilon > 0$,
  - minimizing the computational cost.
- $(Y_h)_{h\in\mathcal{H}} \in \mathbf{L}^2(\mathbf{P})$ a family of simulatable real random variables, approaching Y_0.

Main assumptions

- Bias expansion (weak error assumption): $WE_{\alpha,\bar{R}}$

$$\mathbf{E}[Y_h] - \mathbf{E}[Y_0] = \sum_{k=1}^{\bar{R}} c_k h^{\alpha k} + o(h^{\alpha \bar{R}}),$$

- Quadratic error (strong error assumption): SE_β

$$\left\|Y_h - Y_0\right\|_2^2 = \mathbf{E}\left[\left|Y_h - Y_0\right|^2\right] \leqslant V_1 h^\beta,$$

- Complexity:

The simulation cost of Y_h is $\kappa_h = \dfrac{\kappa}{h}$.

Multilevel Monte Carlo – MLMC (Giles, 08)

The construction of a Multilevel Monte Carlo estimator lies on the two following ideas

- Simulating Y_h is less expensive than simulating $Y_{\frac{h}{M}}$
- $\mathbf{E}\left[Y_{\frac{h}{M}}\right] = \mathbf{E}\left[Y_h\right] + \mathbf{E}\left[Y_{\frac{h}{M}} - Y_h\right]$

Two-level Monte Carlo estimator:

$$\frac{1}{N_1}\sum_{k=1}^{N_1} Y_h^{(1),k} + \frac{1}{N_2}\sum_{k=1}^{N_2}\left(Y_{\frac{h}{M}}^{(2),k} - Y_h^{(2),k}\right)$$

We set $h_j = h/M^{j-1}$.

Definition (Multilevel Monte Carlo estimator)

$$\frac{1}{N_1}\sum_{k=1}^{N_1} Y_h^{(1),k} + \sum_{j=2}^{R}\frac{1}{N_j}\sum_{k=1}^{N_j}\left(Y_{h_j}^{(j),k} - Y_{h_{j-1}}^{(j),k}\right)$$

From Multistep to Multilevel – ML2R (Lemaire, Pagès, 14)

- Take advantage of the whole bias expansion.
- $\mathbf{W}_j^R = \sum_{k=j}^R w_k$, $(w_1, \ldots, w_R)$ solution of a Vandermonde system.

Definition (Multilevel Richardson-Romberg estimator)

$$\frac{1}{N_1}\sum_{k=1}^{N_1} Y_h^{(1),k} + \sum_{j=2}^{R} \frac{\mathbf{W}_j^R}{N_j}\sum_{k=1}^{N_j}\left(Y_{h_j}^{(j),k} - Y_{h_{j-1}}^{(j),k}\right)$$

Cost Minimization

$$I_\pi^N := \frac{1}{N_1}\sum_{k=1}^{N_1} Y_h^{(1),k} + \sum_{j=2}^{R}\frac{\mathbf{W}_j^R}{N_j}\sum_{k=1}^{N_j}\left(Y_{h_j}^{(j),k} - Y_{h_{j-1}}^{(j),k}\right)$$

Parameters (π, N):

- Multilevel estimators: $\pi = (h, R, q = (q_1, \ldots, q_R))$. $N_j = q_j N$.
- Crude Monte Carlo estimator: $\pi = h$.

Optimal parameters $(\pi(\varepsilon), N(\varepsilon))$

Assuming $WE_{\alpha,\bar{R}}$ and SE_β, minimize the simulation cost, given a $\mathbf{L}^2$ – error ε:

$$(\pi(\varepsilon), N(\varepsilon)) = \underset{\left\|I_\pi^N - I_0\right\|_2 \le \varepsilon}{\operatorname{argmin}} \ \mathrm{Cost}(I_\pi^N).$$

Reference without bias (if Y_0 was simulatable): $\mathrm{Cost}(I^N(\varepsilon)) = K\varepsilon^{-2}$

Crude Monte Carlo (Duffie, Glynn)

$$\mathrm{Cost}\left(I_\pi^N(\varepsilon)\right) \precsim K(\alpha)\varepsilon^{-\left(2+\frac{1}{\alpha}\right)}$$

Multilevel (Giles MLMC — Lemaire, Pagès ML2R)

- $\beta > 1$: $\mathrm{Cost}\left(I_\pi^N(\varepsilon)\right) \precsim K(\alpha,\beta,M)\varepsilon^{-2}$.
- $\beta \leq 1$: $\mathrm{Cost}\left(I_\pi^N(\varepsilon)\right) \precsim K(\alpha,\beta,M)v(\varepsilon)$.

	$v_{MLMC}(\varepsilon)$	$v_{ML2R}(\varepsilon)$
$\beta = 1$	$\varepsilon^{-2}\log(1/\varepsilon)^2$	$\varepsilon^{-2}\log(1/\varepsilon)$
$\beta < 1$	$\varepsilon^{-2-\frac{1-\beta}{\alpha}}$	$\varepsilon^{-2}e^{\frac{1-\beta}{\sqrt{\alpha}}\sqrt{2\log(1/\varepsilon)\log(M)}}$

When $\beta > 1$:

Clark Cameron model:

$$\begin{cases} dU_t = S_t dW_t^1, \\ dS_t = \mu dt + dW_t^2 \end{cases}$$

$\mu = 1, T = 1, U_0 = 0, S_0 = 0$

Payoff:

$Y_0 = 10\,\mathbf{E}\left[\cos(U_T)\right]$

True value: 7.14556

Euler: $\alpha = 1$, $\beta = 1$.

Antithetic Giles-Szpruch:
$\alpha = 1$, $\beta = 2 > 1$.

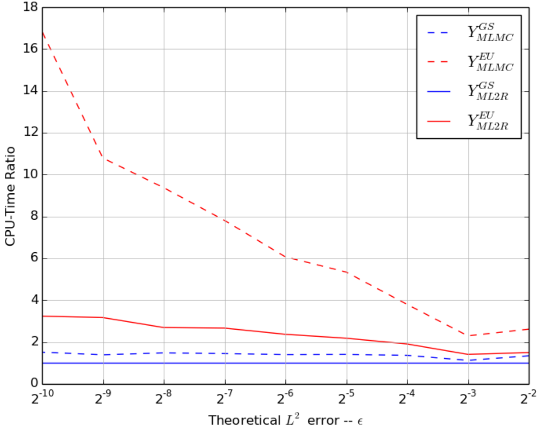

Figure: CPU-Time Ratios

Asymptotic behaviour

- We notice that the L^2 convergence holds by construction, since $\left\|I_\pi^N(\varepsilon) - I_0\right\|_2 \leq \varepsilon$.
- Strong Law of Large Numbers: For all sequence $(\varepsilon_k)_{k\geq 1}$ such that $\sum_{k\geq 1} \varepsilon_k^2 < +\infty$,

$$\sum_{k\geq 1} \mathbf{E}\left[\left|I_\pi^N(\varepsilon_k) - I_0\right|^2\right] < +\infty,$$

  hence $I_\pi^N(\varepsilon_k) \xrightarrow{a.s.} I_0$.
- We can weaken the assumption on the sequence $(\varepsilon_k)_{k\geq 1}$ when Y_h has finite moments of order bigger than 2.

Central Limit Theorem

We write

$$\frac{I_\pi^N(\varepsilon) - I_0}{\varepsilon} = \frac{\mu(h, R(\varepsilon), M)}{\varepsilon} + \frac{1}{\varepsilon\sqrt{N(\varepsilon)}}\sqrt{N(\varepsilon)}\widetilde{I}_\varepsilon^1 + \frac{\widetilde{I}_\varepsilon^2}{\varepsilon},$$

$\mu(h, R(\varepsilon), M)$: bias,

$\widetilde{I}_\varepsilon^1$: first centered coarse level,

$\widetilde{I}_\varepsilon^2$: sum of the corrective centered fine levels.

- $\sqrt{N(\varepsilon)}\widetilde{I}_\varepsilon^1 \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_1^2)$
- $\frac{\widetilde{I}_\varepsilon^2}{\varepsilon} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma_2^2)$

Thank you for your attention

http://simulations.lpma-paris.fr/